NAZARBAYEV
UNIVERSITY

# Study of tuberculosis transmission dynamics using SIR epidemic model with three types of population recruitment rate


By:

**Aigerim Kalizhanova**

Supervisor:

**Ardak Kashkynbayev**
**Associate Professor**

Second Reader:

**Abduzhappar Gaipov**
**Associate Professor**


Submitted to the Department of Mathematics
in partial fulfillment of the
requirements for the degree of
Master of Science in Applied Mathematics

**Nazarbayev University**

**Astana - Kazakhstan**

**May 2024**

# Abstract

This study considers the basic SIR model with three distinct population growth rates to model the tuberculosis transmission dynamics in Kazakhstan. The local stability of disease-free,endemic and extinction equilibrium points of the the models was examined, and the transcritical bifurcation scenario was observed. The effect of growth type on the epidemiological situation in the country was observed through the dynamical analysis and numerical simulations using the data retrieved from the Unified National Electronic Health System (UNEHS). Numerical results reported the basic reproduction number of each of the models to be $R_0 < 1$. Numerical findings support theoretical results. It was revealed that tuberculosis dynamics is approaching its disease-free equilibrium state with $R_0 < 1$, as it was verified by the stability analysis of all three models. SIR model with logistic recruitment rate was proven to give the most favourable results for the epidemic situation in the country.

# Acknowledgments

I would like to express my sincere gratitude to my supervisor, professor Ardak Kashkynbayev, for his guidance and help throughout this thesis work. His patience and will to share the knowledge and experience inspired me during this study.

I also thank Dr.Abduzhappar Gaipov, an Associate Professor at the School of Medicine of Nazarbayev University, for providing his valuable feedback throughout the whole period of this research.

I would like to thank Dr.Sauran Yerdessov, a public health specialist, who provided the data along with assistance during simulations.

I also want to thank my fellow research assistants, Aisha, Kathiresan and Otankhan, who were always ready to answer my questions and give a hand, and inspired me during the whole process of this thesis work.

Last but not least, I would like to thank my family, friends and a significant other for their support and unconditional love.

# Table of Contents

# Chapter One

# Introduction

This thesis is dedicated to mathematical modeling of epidemiological situation in Kazakhstan. The work is based on modeling of the most widespread and contagious disease - tuberculosis (TB). Mathematical modeling can be of great use when analyzing disease transmission dynamics in the community. It is possible to define the rate of transmission and the rate of recovery, and to predict further disease dynamics for a few months or even years depending on the time range of the given data and the model characteristics applied. Disease dynamics are visualized via numerical simulations. Using the results of model analysis and simulations, it is possible to observe the effect of changing model parameter values, hence proposing a control strategy to slow down the epidemics. Moreover, by considering vital dynamics, more realistic scenarios of disease transmission can be observed and analyzed.

The are numerous research works done on the field of mathematical modeling in epidemiology. The very first model was proposed by Daniel Bernoulli in the 18th century to describe the smallpox epidemic [1] . It served as a foundation for seminal studies in mathematical modeling in epidemiology. The ground for research works of the contemporary world was laid by such scientists as Ronald Ross, W.O. Kermack and A. G. McKendrick, giving us the possibility to study such infectious diseases as malaria, plague, tuberculosis, viral meningitis, and more [2, 3]. Recent publications on mathematical modeling, although having the same basis, differ by the level of model complexity in accordance with the aim of the study. The mentioned basis is the SIR model proposed by Kermack and McKendrick [3], where each letter in the abbreviation indicate the group of the target population: S - susceptible people, I - infected people and R - recovered people. Chen et al., Kereyu and Demie, Das et al. studied tuberculosis dynamics using modified SIR and aimed at age-specific groups, and control strategies such as media awareness, distancing and treatment [4, 5, 6]. Chen et al., whose study focused on the interaction between students and the non-student part of the Chinese population, revealed that the high prevalence in a non-student population greatly influences students' health. Hence it was proposed to perform TB screening among students [4]. Whereas Kereyu and Demie, as the result of numerical comparison of three different control strategies: distancing, case finding, and

treatment efforts, were able to reveal that the best strategy is the combination of the three [5]. Das et al. focused on the media awareness strategy and confirmed that as the awareness level is increased, TB prevalence gradually decreases [6]. There are also works that compare the basic SIR with its compartment modified versions. For instance, Ucakan et al. considered adding more compartments to the model and analyzed their ability to simulate the dynamics of TB [7]. The models are SIR, SEIR, BSEIR. Here, E is the compartment of exposed people, meaning the group of infected people but with no visible symptoms of TB, while B stands for the group of Bacillus Calmette-Guérin (BCG) vaccinated people. BCG is the vaccine used to prevent TB in countries with high prevalence of the infection, including Kazakhstan [8]. Comparing the three models, authors came to the conclusion that all of them give similar results in terms of dynamics simulation, calculation of the infection and recovery rates, and the number of secondary infections. Hence, if no specific parameter is needed for the study, SIR itself performs well [7].

In this work, we will study the SIR model incorporating three different population recruitment rates and observe the effect of growth rate on the ability of the disease to spread, by calculating the basic reproduction number of each model and evaluating the stability of endemic and disease-free states. It will allow us to understand how population growth influences the TB dynamics, thus giving the opportunity to think about new disease control strategies.

## 1.1 Motivation

In the modern age, extensive studies are being done on TB using various tools: mathematical and statistical models [6, 5], comprehensive data analysis [9] and quantitative analysis via questionnaires [10]. Although the last two tools are widely used in Kazakhstan, mathematical models are rarely employed.

Tuberculosis (TB) is caused by Mycobacterium tuberculosis bacteria. It is an air-borne disease, which spreads when an infected person coughs, sneezes or spits [11]. Common symptoms of a disease include chest pain, weakness, fatigue, fever, cough and weight loss. Even though it is a curable disease, it also gets resistant to drugs [11]. Hence humanity must not stop looking

for ways to minimize the spread of the infection.

According to a recent comprehensive study on TB in Kazakhstan from 2014 to 2019, the incidence rate fell from 227.0 to 69.1 cases per 100 000 population, while the disease mortality rate increased for more than 1000 people [12]. TB was the second leading infectious killer in the World, after the coronavirus in 2022 [13]. Moreover, due to the COVID-19 pandemic and the restrictions during the quarantine period, patients with TB did not have the same opportunities as before to get treatment. This resulted in high mortality rates among TB patients in the pandemic period [13]. In 2023 the World Health Organization (WHO) reported the latest analysis of global TB dynamics and introduced three lists of high burden countries (HBC). They were established for TB, HIV-associated TB and drug-resistant TB, and the countries are under the control of the organization. Only the top 20 countries with the highest incidence rates and 10 countries with the most severe cases that are not in the top 20, but meet the minimum criteria for absolute number of incidence rates are considered in these lists. Kazakhstan is included in two of the lists: TB and drug-resistant TB as a country that does not meet the thresholds: maximum of 10 000 new cases per year for TB and 1000 incidences for HIV-associated TB and drug-resistant TB [13]. Therefore, studying the dynamics of these infectious diseases might set the stage for further research using mathematical modeling.

TB modeling can be done using various methods depending on the aim of the study. As described in the previous section, there are models considering various compartments, control strategies and parameters. This study will focus on the model that considers three different recruitment rates. Such study was done by Zhao et al. to describe the dynamics of HIV (human immunodeficiency virus) along with HIV preventive measures. As a result, it was proved that the stability of equilibrium points and the HIV preventive measures' impact strongly depend on the recruitment mechanism [14]. This study aims at investigating the effect of recruitment rates on the TB dynamics in Kazakhstan and examine the stability of equilibrium points.

## 1.2 Guide to Thesis

This thesis work is structured in the way described below.

**Chapter Two** gives some background on epidemiological modeling. The basic SIR model with and without vital dynamics is discussed along with its dynamical analysis. One of the most important parameters - the basic reproduction number $R_0$ is introduced. Moreover, tuberculosis (TB) modeling history and the current trends of mathematical modeling of TB are discussed.

**Chapter Three** introduces differential equation theory and techniques employed for model formulation and its analysis. To be more precise, the general form of the system of nonlinear ordinary first-order differential equations, its existence and uniqueness theorems, local stability principles and bifurcation phenomenon are introduced.

**Chapter Four** formulates the mathematical model for TB transmission, finds equilibrium points of the models, performs stability analysis, identifies bifurcation scenarios and summarizes the results into new theorems.

**Chapter Five** performs numerical simulations of the models using the data on monthly TB cases in Kazakhstan, and compares the results with theoretical findings. In addition, the performance of the models is evaluated using well-known statistical metrics such as R-squared, RMSE, MAE and MAPE.

**Chapter Six** discusses the results obtained in chapters four and five, and examines whether the objective of this thesis work is achieved. Moreover, the limitations of this work are reviewed.

**Chapter Seven** concludes the work and summarizes the findings of this study, and proposes new ideas for further research.

# Chapter Two

# Epidemic Modeling Background

## 2.1 Kermack-McKendrick Model

### 2.1.1 The Basic SIR Model

Epidemiological processes are often represented as mathematical models for further analysis. The most basic model was proposed by Kermack and McKendrick in their phenomenal work series "Contributions to the Mathematical Theory of Epidemics" [15, 16, 17], and has the following form:

$$\begin{cases} S'(t) = -\beta S(t)I(t), \\ I'(t) = \beta S(t)I(t) - \gamma I(t), \\ R'(t) = \gamma I(t). \end{cases} \tag{2.1}$$

The idea here is to divide the host population of size $N$ into three compartments: susceptible ($S$ – people that are likely to get infected, no immunity), infected ($I$ – infected individuals who can infect susceptible ones in case of contact) and recovered ($R$ – successfully recovered, hence immune people). The total population size is then defined as $N = S + I + R$ and is held constant. The flow of people over time between compartments is described using the system of ordinary differential equations (2.1). As this system is solved, we get sizes of compartments: $S(t)$, $I(t)$, $R(t)$ at time $t$. To comply with biological norms, the number of people in each compartment must be non-negative. Since $N$ can be sufficiently large, we may consider $S$, $I$ and $R$ to be continuous functions of time and represent rates of changes in them as in (2.1) [18]. The interaction between a susceptible person and an infectious person is described by a mass-action law, $\beta SI$, where $\beta$ is the rate at which one gets infected after a contact with an infectious person. The rate at which infected people recover from the disease is $\gamma$.

This model was later modified by ReVelle, who proposed to use a probabilistic infection rate $\frac{\beta SI}{N}$. He proved using this approach that the infection rate linearly depends on the prevalence. The model now looks like the system (2.2) which is mostly employed in studies nowadays. He

stated that $\frac{\beta SI}{N}$ is a more appropriate infection rate for a growing population and used it in the first nonlinear system to model the tuberculosis [19].

$$\begin{cases} S'(t) = -\frac{\beta S(t)I(t)}{N(t)}, \\ I'(t) = \frac{\beta S(t)I(t)}{N(t)} - \gamma I(t), \\ R'(t) = \gamma I(t). \end{cases} \tag{2.2}$$

This model's simplicity allows to find an exact solution of I(t) [18]:

$$I(t) = I(0) + S(0) - S(t) + \frac{1}{R_0} ln[\frac{S(t)}{S(0)}] \tag{2.3}$$

As it can be observed, the solution depends on another variable S, hence making it impossible to solve the system analytically. However, it is possible to do it numerically. Earn D.J. employed an Eurler's method [18] and considered the epidemics of measles in New York city in 1962, while Azizan et al. used Runge-Kutta method to explore the dynamics of tuberculosis in Malaysia in 2021 [20]. Both methods can easily be implemented in any programming language.

### 2.1.2 The Basic Reproduction Number and Stability

The parameter $R_0$ presented in equation (2.3) defines the average number of secondary infections caused by a single infected person and is called the basic reproduction number. It is a threshold quantity and is significant in model analysis due to its ability to define the further dynamics of the disease. If $R_0 > 1$, the disease becomes endemic and if $R_0 < 1$, the population is reaching the disease-free equilibrium, meaning the declining trend in the disease dynamics [21].

The formula for $R_0$ depends on the model. Considering the model (2.1), we observe that when a newly infected person enters a fully susceptible population ( $S(0) = N$), he/she may infect others at the rate $\beta N$ during their infectious period $1/\gamma$. Hence, on average, an infectious person may infect $R_0 = \frac{\beta N}{\gamma}$ people [18]. There is another method for calculating $R_0$, which

is presented in the paper by van den Driessche and Watmough [22]. Authors take the spectral radius (i.e. the largest eigenvalue) of a next generation matrix as the basic reproduction number. Same technique was used by van den Driessche to calculate $R_0$ for such diseases as West Nile virus in birds, anthrax in animals and Zika in humans; and to propose control strategies based on the results [21]. This method suits best for more complex models with several compartments of infected people than the basic SIR, although still applicable to the latter model. As van den Driessche noted, $R_0$ is simply the product of transmission rate and the average infectious period [22].

Moreover, the basic reproduction number plays a great role in stability analysis of the disease dynamics. Since the system we are discussing consists of differential equations, it is possible to find their roots by setting the right hand-sides of equations equal to zero.

$$\begin{cases} -\beta S(t)I(t) = 0, \\ \beta S(t)I(t) - \gamma I(t) = 0. \end{cases}$$

We usually do not consider the last line of the system (2.1), since it solely depends on the variable $I$. By solving this system for $S$ and $I$, we get to the roots of the system. These roots are called equilibrium points $(S^*, I^*, R^*)$. When studying infection transmission, researchers are interested where the disease dynamics is heading to; whether it is a disease-free equilibrium points or an endemic equilibrium point. The first case is favorable since it means that the disease dies out at some point of time, while the dynamics towards the latter equilibrium means that the population of interest is in danger and the number of disease cases will not reduce. To learn this, stability analysis of these equilibria is done. In the case of model (2.2), there are only two equilibria: $(0,0,0)$ and $(\frac{\gamma}{\beta},0,0)$. The first one is the trivial point, indicating there is no population, hence, no infection. Second equilibrium can give more information on the dynamics. As it can seen, if $S^* = \frac{\gamma}{\beta}$, then there is no infection in the population: $I = 0$, $I' = 0$. Let us consider the second line of the system (2.1) and rearrange it (omitting the time variable

$t$ for the sake of notation):

$$I' = (\beta S - \gamma)I$$

Looking at this equation, we observe that if $S > \frac{\gamma}{\beta}$, then $I' > 0$ and $S' < 0$. Let us consider the dynamics from the very beginning, given the initial condition $S(0) = \frac{\gamma}{\beta}$. If $S(0) > \frac{\gamma}{\beta}$, then the number of infected people might reach its maximum before the decline, while otherwise, $I$ goes to 0, meaning no further epidemics. By rearranging this condition, we get the expression for basic reproduction number:

$$R_0 = \frac{\beta S(0)}{\gamma} \tag{2.4}$$

Now, we can discuss the stability of disease dynamics in terms of $R_0$. If $R_0 < 1$ then there is no epidemic since the number of infected people goes to zero; while if $R_0 > 1$ then the number of infected people reaches its maximum point at $S = \frac{\gamma}{\beta}$ and reduces to zero afterwards.

As we can see from the above results, equilibrium points exchange the stability at the threshold value $R_0 = 1$. While the disease free state is stable for $R_0 < 1$, the endemic case is unstable; and vice versa. This process is called bifurcation [23]. It is important to know the bifurcation point in disease transmission analysis, since it is the critical point, that should or should not be passed to keep the stability of one or another equilibrium point. In this case, a bifurcation point is $R_0 = 1$, meaning that if the number of secondary infections exceeds 1, then the disease free equilibrium loses the stability, and the population moves towards the endemic situation. Therefore, we want to keep this parameter below 1 by adding various disease control strategies[14, 7, 24].

## 2.2 SIR with Vital Dynamics

The basic model described in the previous section is not able to fully describe the dynamics of an infectious disease in the real life. Although the basic SIR model offers a straightforward representation of disease dynamics with only three compartments and ignores demography and control measures, which could be beneficial for accuracy of simulations and analyses, it is able

to give the general disease trends and its dependence on model parameters. However, since no vital dynamics were incorporated, the SIR model can only perform well for no more than 2-3 year period dynamics[25, 26]. Moreover, SIR with vital dynamics proves to be better at future disease dynamics prediction [20]. Therefore, if there is a need to analyse or predict the disease dynamics, one might want to consider adding vital dynamics.

The very first SIR model with vital dynamics was proposed by H.E.Soper in 1929, who added constant birth rate into the susceptible class and the death rate into the recovered class, and labeled as $\mu$ [27]:

$$\begin{cases} S'(t) = \mu K - \beta SI, \\ I'(t) = \beta SI - \gamma I, \\ R'(t) = \gamma I - \mu R. \end{cases} \tag{2.5}$$

However, this model is considered unacceptable both mathematically and biologically [26]. Brauer explained that connection of birth of susceptibles and the death of only recovered people does not make sense from the biological point of view. In addition, if both initial values of $I(t)$ and $R(t)$ are sufficiently small, then the number of recovered people might get negative at some point of time during the period of interest [26]. This contradicts well-posedness conditions of the whole problem: the number of people in each compartment is non-negative throughout the whole period.

In the modern age, we rely on the following general form of SIR with vital dynamics such as natural birth and death rates:

$$\begin{cases} S'(t) = \Lambda(N) - \beta(N)SI - \mu S, \\ I'(t) = \beta(N)SI - \gamma I - \mu I, \\ R'(t) = \gamma I - \mu R, \\ N'(t) = \Lambda(N) - \mu N. \end{cases} \tag{2.6}$$

Here, $\Lambda(N)$ represents the recruitment rate into the susceptible class. We can choose it to be

constant or greater than 0. $\mu$ is the natural death rate. Since the population size is non-constant anymore, it has its own rate of change. In addition, it might reach its limiting capacity $K = \Lambda/\mu$, called the carrying capacity or the maximum possible population size [26].

To analyze this model, we start by examining whether it is well-posed. We observe that if $S = 0, I = 0$ and $R = 0$, then $S' \geq 0$, $I' \geq 0$ and $R' \geq 0$, respectively. Moreover, if we assume that initially $N$ has reached its carrying capacity $N = K$, then $N' = \Lambda(K) - \mu K = 0$. Hence, $N \leq K$ for $t \geq 0$. Which confirms that the number of people in compartments cannot be negative if we begin with $S > 0$, $I > 0$ and $0 < N \leq K$.

As the well-posedness is confirmed, equilibrium points of the system are searched for. We begin by setting each of the equations of the system equal to 0.

$$\begin{cases} \Lambda(N) - \beta(N)SI - \mu S = 0, \\ \beta(N)SI - \gamma I - \mu I = 0, \\ \gamma I - \mu R = 0. \end{cases} \tag{2.7}$$

As the result, we get two different equilibria: disease free equilibrium (DFE) and the endemic equilibrium (EE).

DFE:

$$(S_0, I_0, R_0) = \left(\frac{\Lambda(N)}{\mu}, 0, 0\right) = (K, 0, 0)$$

EE:

$$(S_\infty, I_\infty, R_\infty) = \left(\frac{\gamma+\mu}{\beta}, \frac{\Lambda(N)}{\gamma+\mu} - \frac{\mu}{\beta}, \frac{\gamma\Lambda(N)}{\gamma+\mu} - \frac{\gamma}{\beta}\right) = \left(\frac{\gamma+\mu}{\beta}, \frac{\mu K}{\gamma+\mu} - \frac{\mu}{\beta}, \frac{\gamma\mu K}{\gamma+\mu} - \frac{\gamma}{\beta}\right)$$

If the system was two dimensional, the system would be linearized around these equilibrium points. Coefficient matrix of the linearized system would give us the eigenvalues, which would indicate stability properties of these equilibrium points [26]. This technique is not suitable for the system (3), hence we apply the Routh-Hurwitz stability criterion to analyze the stability.

This technique is discussed in the following section.

## 2.3 Modeling Tuberculosis Mathematically

Tuberculosis (TB) is one of the most contagious diseases. Humanity has been fighting it for centuries. The first modeling approach (not necessarily mathematical) was applied by Wade H. Frost in 1937, who mainly discussed a TB preventive measure such as avoiding getting infected with tubercle bacillus, which is the bacteria causing the disease. Frost predicted that in this case, the number of incidences in the United States would decline [28]. This prediction was verified by Floyd M. Feldmann in 1957, who used the official data on TB cases and prepared visual representation of the disease dynamics [29]. A year after, Palmer et al. presented the study summarizing the two controlled BCG trials in Puerto-Rico and Muscogee and Russell counties of the United States. As the last points of their study, Palmer et al. pointed out that the vaccine is only efficient for those who are not infected, but will be and the vaccine might not be very useful in controlling the TB in the country. Hence, they took the position stating that "in most situations in this country today the advantages of vaccination are outweighed by the disadvantages" [30].

The process of modeling TB mathematically began in 1962 by Waaler et al. They split the population into three compartments: noninfected (suscpetibles) , and infected noncases (latent cases) and cases (infectious). Authors also consider vital dynamics: number of births and deaths, and the number of healed patients. The model consists of five equations in total, describing the flow between the compartments employing equations constructed with subtraction and addition of the number of people in each compartment and the numbers indicating birth, death and recovery cases. The study provided the three examples of how the model could be used: to fit the model to data and illustrate the time trend of the disease dynamics, to show the effect of a control strategy, and to observe the effect of a BCG vaccine on the disease trends [31]. This model was further developed by such scientists as Brogger and Revelle, who modified the contact rate between an infectious person and a susceptible person in terms of statistical rules

[32, 33].

Nowadays researchers take this model as the basis and try to modify it depending on the up-to-date processes happening in certain epidemiological situation and the goals of the study. First, we might consider studies that divide the population into more than three basic compartments. Bilal et al.(2021) studied two more groups of people: recovered and resistant. Using the model with five compartments, authors examined the effect of two non-linear controllers designed for treatment and vaccination of infected people [34]. Whereas Zhao et al. (2017) studied the effect of age on the TB transmission in China by exploiting the SEIR (Suscepltible-Exposed-Infected-Recovered) model with age groupings, considering three age categories: children, the middle-ages and senior. The study revealed that different age groups affect the dynamics differently. For instance, the BCG vaccine is more efficient for children, while has almost no effect on older generation. However, since the older people are at high risk of disease induced morbidity, there is a ground to think of another vaccine that might suit this age group [35]. A similar study was done in South Korea in 2021 by Lee et al., who studied two different age groups: elderly and non-elderly. The study showed that TB treatment intensity must be increased for elderly patients, while the importance of early disease detection and treatment of latently infected patients was proven to be equal for both groups [36].

Aside from model structure, model parameters may also be varied. It is a common practice to take into an account birth and death rates to model the TB transmission dynamics. Recent works on this topic mostly consider either constant $\Lambda$ [24, 37] or proportional $rN$ [4, 6, 34, 38] recruitment rates, where $\Lambda$ is the fixed number of newborns or immigrants at a certain time period (day/month/year) and $r$ is the proportional recruitment, meaning the rate at which new people join the community. However, there are few considering logistic growth of the population. The reason might be the case when the population indeed grows linearly over a longer period of time. In addition, when the complexity of the model to be analyzed is increased by other factors, as discussed above, it might be better to use these two growth rates to simplify the analysis. Despite these factors, there is a study proving the importance of the population

recruitment rate in the discussion of the stability of disease free and endemic equilibrium states and the bifurcation analysis. It was revealed that the three growth rates affect differently the model and may cause more that 6 percentage point variation in disease prevalence over the period of 40 years [14]. The study focused on the HIV epidemics and its prevention, while this paper's aim is to model tuberculosis dynamics in Kazakhstan and compare the effect of various recruitment rates from the mathematical point of view, by considering simulations, stability and bifurcation analysis.

# Chapter Three

# Differential Equation Theory

This chapter focuses on some of the existence, uniqueness and stability theorems for ordinary differential equations that are applicable to mathematical modeling in epidemiology. Let us consider the general form of an ordinary differential equation (ODE):

$$x'(t) = f(x,t) \tag{3.1}$$

which is a non-autonomous equation due to the dependence on the time variable $t$ and the state variable $x(t)$. In epidemic modeling, we usually consider initial conditions of the following form:

$$x(t_0) = x_0 \tag{3.2}$$

In this work, the system of such differential equations is employed, as it can be seen from systems (1-3). However, in basic epidemic models such as SIR (1-3), the system does not explicitly depend on time $t$, hence the systems considered in this work are autonomous. The state variables indicate the number of susceptible ($S$), infected ($I$) and recovered ($R$) people and their continuous time dependence, while initial conditions represent the size of each compartment at time $t_0$. The general form of the autonomous system of first order nonlinear ordinary differential equations is considered to be

$$\begin{cases} x_1'(t) &= f_1(x_1,x_2,...,x_n), \\ ... & \\ x_n'(t) &= f_n(x_1,x_2,...,x_n). \end{cases} \tag{3.3}$$

This system can be written compactly in the following form:

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) \tag{3.4}$$

where $\mathbf{x}(t) = (x_1(t), ..., x_n(t))^T$ and

$$\mathbf{F}(\mathbf{x}) = (f_1(x_1, x_2, ..., x_n), ..., f_n(x_1, x_2, ..., x_n))^T$$

, and the general initial condition is $\mathbf{x}(t_0) = \mathbf{x}_0$, where $\mathbf{x}_0 = (x_1(t_0), x_2(t_0), ..., x_n(t_0))^T$. The solution of such a system has to satisfy it on a given interval and has to be defined and continuously differentiable on the given time interval. Although solutions of such systems exist in most of the cases, it might be challenging to find them [39]. Hence, we first verify that the solution exists and is unique given specific initial conditions.

## 3.1 Existence and Uniqueness

To ensure that the system has the unique solution, we rely on the Fundamental theorem of ordinary differential equations [40].

**Theorem 3.1.1 (Existence and Uniqueness)** *Consider the initial value problem (IVP)*

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{F}(\boldsymbol{x})$$

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0$$

*where $\boldsymbol{x}_0 \in \mathbb{R}^n$, and suppose that $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable. Then,*

- *there exists a solution of this IVP;*
- *the solution is unique.*

*In other words, there exists a constant $a > 0$ and a unique solution*

$$\boldsymbol{x} : (t_0 - a, t_0 + a) \to \mathbb{R}^n$$

*that satisfies the initial condition.*

To ensure that the models studied in this work do have unique solutions, it is enough to show that the solutions exist in a compact set $\{(S,I,R) : S+I+R=N, S \geq 0, I \geq 0, R \geq 0\}$. This was illustrated in a previous chapter.

## 3.2 Local Stability

Once the existence and uniqueness of the solution is verified, we focus on steady states of the system and their stability properties. Steady states $\mathbf{x}^*$ are also called equilibrium points/solutions, and they are obtained as a result of the following system:

$$\mathbf{F}(\mathbf{x}^*) = 0$$

P.J.Olver and C.Shakiban proved an important proposition explaining why we look at equilibrium points when performing stability analysis. He formulated that into the following proposition.

**Proposition 3.2.1** *Let $x(t)$ be a solution to the system (3.4), with $\boldsymbol{F}$ being continuously differentiable, such that $\lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{x}^*$. Then $\boldsymbol{x}^*$ is an equilibrium solution and $\boldsymbol{F}(\boldsymbol{x}^*) = 0$.*

In other words, if a solution of an autonomous system like in (3) converges to the limit point, then that point is an equilibrium point. The proof can be found in [39]. This point is called stable if all other solutions approach it as time passes. It is also important to distinguish that, if every solution starting near the equilibrium point tends to it over time, then the equilibrium point is asymptotically stable; while if the solution starting near the point stays around it, then it a stable point.

For the first order differential equations, we usually rely on monotonicity of the solution to examine its stability. However, for higher order systems (in our case it is 3D) we perform linearization to do stability analysis. Taking the derivative $f'(x^*)$ of the right hand-side of the differential equation gives the linear approximation of the function $f(x)$ near the equilibrium point by giving the slope of a tangent line. We do the analogue of this technique for a higher

dimensional system (3). In this case, $\mathbf{F}'(\mathbf{x}^*)$ is an $n \times n$ Jacobian matrix $J(\mathbf{x}^*)$ evaluated at equilibria. As in the one dimensional case, we use it to approximate the function $\mathbf{F}(\mathbf{x})$ near the equilibria, and write out its Taylor expansion:

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}^*) + \mathbf{F}'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*).$$

Keeping in mind that the function at an equilibrium point is zero:

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*). \tag{3.5}$$

Denoting the difference $\mathbf{v} = \mathbf{x} - \mathbf{x}^*$, we get the linearized system:

$$\frac{d\mathbf{v}}{dt} = J(\mathbf{x}^*)\mathbf{v} \tag{3.6}$$

To examine the stability of equilibrium points, we may rely on the following theorem [39]

**Theorem 3.2.1** *Let $\boldsymbol{x}^*$ be an equilibrium point for the system (3.4). If all eigenvalues $\lambda$ of the Jacobian matrix $J = \boldsymbol{F}'(\boldsymbol{x}^*)$ have negative real parts, then $\boldsymbol{x}^*$ is said to be asymptotically stable. Otherwise, $\boldsymbol{x}^*$ is unstable.*

However, some systems linearized around the equilibria might not give eigenvalues that are straightforward to analyze. Such a system is described in chapter four. In this case, the stability of the equilibrium point is examined using the characteristic equation obtained as a result of

$$det(J - \lambda I) = 0$$

and the Routh-Hurwitz criterion. It states that given the equation of the form

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$$

the necessary and sufficient conditions for all roots to have negative real parts are

$$a_1 > 0, \qquad\qquad a_1 a_2 > a_3 > 0.$$

Hence, it is necessary to verify that these conditions are satisfied, before stating the stability property of an equilibrium point. As the stability results are ready, it is a common practice to observe that at some values of a certain parameter, the stability of the solutions of the system change. This process is called bifurcation.

#### 3.2.0.1 Bifurcation

This section is dedicated to the discussion of bifurcations, a significant process in the analysis of nonlinear differential systems. Let us consider the general formulation of a system of nonlinear differential equations:

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}_a(\mathbf{x}) = \mathbf{F}(\mathbf{x}, a), \tag{3.7}$$

where $a \in \mathbb{R}$ is a parameter and $\mathbf{F}_a : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is assumed to depend on $a$ in an infinitely differentiable manner. We say that there a bifurcation if the structure of the system solutions change significantly as the parameter $a$ is varied. A critical point $a_0$ when the behavior of solutions change is called a bifurcation point. Let us recall the epidemiological system in chapter 2 and the analysis done in section 2.1.2. The bifurcation point is $R_0 = 1$, and depending on whether it gets greater or less that 1, the stability of equilibrium points change. If $R_0 < 1$, all solutions of the system approach the disease free equilibrium point, otherwise they approach the endemic equilibrium.

Consider a general first-order differential equation

$$\frac{dx}{dt} = f_a(x). \tag{3.8}$$

If $x_0$ is an equilibrium point of this equation then $f_a(x_0) = 0$. If it is not, then any small changes to the parameter $a$ are not able to alter the behavior of the solution near this equilibrium point. Hence, bifurcations in first-order differential equations only occur when $f'_a(x_0) = 0$, i.e. in a non-hyperbolic case [40].

There are various types of bifurcation such as transcritical, saddle-node, pitchfork bifurcations. Saddle-node bifurcation is a common example of bifurcation. It happens when at first, there are two equilibrium points with opposite stabilities. As the parameter reaches the critical point, equilibrium points collapse. On the bifurcation diagram, this process looks like a parabola. It generally occurs when

$$\mathbf{F}(0,0) = \mathbf{F_x}(0,0) = 0,$$
$$\mathbf{F}_a(0,0) \neq 0,$$
$$\mathbf{F_{xx}}(0,0) \neq 0,$$

where the subscript means the partial derivative of the function in terms of $a$ and $x$ [41].

Transcritical bifurcation also requires two equilibrium points, which exchange stabilities at $a = 0$. In general, this type of bifurcation occurs when

$$\mathbf{F}(0,0) = \mathbf{F_x}(0,0) = \mathbf{F}_a(0,0) = 0,$$

and $\mathbf{F_{xx}}(0,0) \neq 0$.

Another type of local bifurcations is a pitchfork bifurcation, and it has two types: supercritical and subcritical. Supercritical pitchfork bifurcation is described by the transition of the system from only one equilibrium point to three points as the parameter value exceeds $a = 0$. While the subcritical pitchfork bifurcation is characterised by the transition from three to one fixed point at the same condition. General conditions for such bifurcation to occur are the

followings[41]:

$$\mathbf{F}(0,0) = \mathbf{F}_{\mathbf{x}}(0,0) = \mathbf{F}_{a}(0,0) = \mathbf{F}_{\mathbf{xx}}(0,0) = 0,$$

$$\mathbf{F}_{\mathbf{x}a}(0,0) \neq 0,$$

$$\mathbf{F}_{\mathbf{xxx}}(0,0) \neq 0.$$

In this work, all of the described techniques are applied to analyze the three models and define conditions for the stability of equilibrium points along with examination of possible bifurcation scenarios.

# Chapter Four

# Mathematical Model for Tuberulosis Transmission

In this chapter we formulate the models to describe the dynamics of TB transmission in Kazakhstan using three different population recruitment rates, analyze the models mathematically including stability analysis and bifurcations. We begin with SIR with constant recruitment rate, continue with proportional, and finish with logistic rate. The models are based on the basic SIR model discussed in chapter 2, the only difference is in the rate of people inflow to the susceptible compartment. Hence, the well-posedness statements are the same for all three models. In general, the model looks as follows:

$$\begin{cases} S'(t) = f(N) - \frac{\beta SI}{N} - \mu S, \\ I'(t) = \frac{\beta SI}{N} - (\gamma + \mu)I, \\ R'(t) = \gamma I - \mu R, \end{cases} \tag{4.1}$$

where $f(N)$ represents the recruitment rate, and will be replaced by the respective type. Parameters of this model are: $\beta$ - the contact rate, $\gamma$ - recovery rate, $\mu$ - natural death rate. The basic reproduction number is

$$R_0 = \frac{\beta}{\mu + \gamma} \tag{4.2}$$

For this model, we assume that initially compartments of susceptible, recovered and infected people are non-empty, while the newly coming individuals automatically get susceptible, i.e. are at risk of getting infected with tuberculosis.

Since we want to solve the systems in future, we assume that we are given initial conditions at $t = 0$: $(S(0), I(0), R(0), N(0))$. In addition, it can be observed that the right hand-side equations of the system (4.1) are continuously differentiable in terms of state variables $S, I, R$. By the existence and uniqueness theorem 3.1.1 presented in chapter 3, we claim that this system has a unique solution given the initial conditions. Moreover, in section 2.2, it was proved that the solution to this system stays in the biologically liable region $\{(S,I,R) : S+I+R = N, S \geq 0, I \geq 0, R \geq 0\}$.

## 4.1 SIR with Constant Recruitment Rate

### 4.1.1 Model Formulation

Let us first formulate the model with constant recruitment rate $\Lambda$ and the probabilistic contact rate $\frac{\beta SI}{N}$:

$$\begin{cases} S'(t) = \Lambda - \frac{\beta SI}{N} - \mu S, \\ I'(t) = \frac{\beta SI}{N} - (\gamma + \mu)I, \\ R'(t) = \gamma I - \mu R. \end{cases} \tag{4.3}$$

The model here considers a constant recruitment rate to the population of size $N = S + I + R$. The rate at which the population size changes is given by

$$N'(t) = \Lambda - \mu N.$$

Since we need the population size to grow, we require $\Lambda - \mu N \geq 0$, which leads to the condition $N \leq \frac{\Lambda}{\mu}$. Hence, the solution we are interested in lies in the region $\{S \geq 0, I \geq 0, R \geq 0, S+I+R \leq \frac{\Lambda}{\mu}\}$. Next step of the analysis is to identify equilibrium points of the system.

### 4.1.2 Equilibria

To find the steady states of the model (4.3), we set each of the equation equal to 0 and remember that the population size is the sum of the three compartments.

$$\begin{cases} S'(t) = \Lambda - \frac{\beta SI}{N} - \mu S = 0, \\ I'(t) = \frac{\beta SI}{N} - (\gamma + \mu)I = 0, \\ R'(t) = \gamma I - \mu R. \end{cases} \tag{4.4}$$

It is also important to understand that there is a carrying capacity of the population, i.e. the maximum population size $K$, which is $K = \frac{\Lambda}{\mu}$. Hence

$$\lim_{t\to\infty} N(t) = \frac{\Lambda}{\mu}.$$

By the theory of asymptotically autonomous systems, if the limit of N is constant, then the system having the limit of N instead of itself is equivalent to the primer system[26].

Let's now consider the second equation and find its roots. First one is $I_0 = 0$ and the second one is $S_\infty = \frac{\mu+\gamma}{\beta}N$. By inserting $I_0$ into the first equation, we can obtain the next point of the disease free equilibrium (DFE): $S_0 = \frac{\Lambda}{\mu}$ and the value of $R$ follows: $R_0 = 0$. Now we can insert the value of the second root into the first equation of the model and obtain the points of an endemic equilibrium (EE). Keeping in mind that $N = \frac{\Lambda}{\mu}$, we find the remaining points. In summary, there are two equilibrium points:

- DFE:

$$(S_0, I_0, R_0) = \left(\frac{\Lambda}{\mu}, 0, 0\right),$$

- EE:

$$(S_\infty, I_\infty, R_\infty) = \left(\frac{\Lambda}{\mu R_0}, (R_0-1)\frac{\Lambda}{\beta}, (R_0-1)\frac{\gamma\Lambda}{\mu\beta}\right).$$

In order to see how the system solutions behave, we want to analyze the stability of these equilibria.

### 4.1.3 Stability

To evaluate the stability of these points, coefficients of a linearization of the system are stored in a Jacobian matrix:

$$J(S,I,R) = \begin{bmatrix} \frac{\partial S'}{\partial S} & \frac{\partial S'}{\partial I} & \frac{\partial S'}{\partial R} \\ \frac{\partial I'}{\partial S} & \frac{\partial I'}{\partial I} & \frac{\partial I'}{\partial R} \\ \frac{\partial R'}{\partial S} & \frac{\partial R'}{\partial I} & \frac{\partial R'}{\partial R} \end{bmatrix}.$$

We now evaluate the matrix at equilibrium points. First, consider the DFE:

$$J(S_0, I_0, R_0) = \begin{bmatrix} -\mu & -\beta & 0 \\ 0 & \beta - \mu - \gamma & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

The eigenvalues of this system are found as the following equation is solved:

$$det(J - \lambda I) = 0.$$

As a result, we get the following eigenvalues:

$$\lambda_1 = -\mu, \qquad \lambda_2 = -\mu, \qquad \lambda_3 = \beta - \mu - \gamma.$$

It can be observed that if $\mu > 0$ and $\frac{\beta}{\gamma+\mu} < 1$, the DFE is stable, while otherwise the equilibrium becomes unstable.

Performing the same steps for an endemic equilibrium, the following Jacobian matrix is obtained:

$$J(S_0, I_0, R_0) = \begin{bmatrix} -\mu R_0 & \frac{\beta}{R_0} & 0 \\ \mu(R_0 - 1) & 0 & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

The corresponding characteristic equation is

$$(-\mu R_0 - \lambda)(\mu + \lambda)\lambda + \frac{1}{R_0}\beta\mu(R_0 - 1)(-\mu - \lambda) = 0. \tag{4.5}$$

It is a cubic polynomial with three possible roots. The Routh-Hurwitz condition can be applied to the analysis[26]. In our case, the equation can be reformulated into

$$\lambda^3 + \mu(R_0 + 1)\lambda^2 + (\mu^2 R_0 + \frac{1}{R_0}\beta\mu(R_0 - 1))\lambda + \frac{1}{R_0}\beta\mu^2(R_0 - 1) = 0, \tag{4.6}$$

such that

$$
\begin{aligned}
a_1 &= \mu(R_0+1),\\
a_2 &= \mu^2 R_0 + \frac{1}{R_0}\beta\mu(R_0-1),\\
a_3 &= \frac{1}{R_0}\beta\mu^2(R_0-1).
\end{aligned}
$$

For the endemic equilibrium to be stable, the Routh-Hurwitz conditions should be satisfied. Which is possible if $a_1 > 0$, meaning that for the inequality $\mu(R_0+1) > 0$ to hold true, we need the natural death rate $\mu$ and the basic reproduction number $R_0$ to be positive. In addition, we know from the analysis of DFE, that for instability of DFE, we need the basic reproduction number to be greater than 1. Now, we check the next condition $a_1a_2 > a_3 > 0$. For that, we first examine if the third coefficient $a_3$ is positive:

$$\frac{1}{R_0}\beta\mu^2(R_0-1) > 0.$$

It can be observed that, indeed, we need to have $R_0 > 1$, and the parameters of death and contact rates positive to satisfy this condition. Lastly, we check if $a_1a_2 > a_3$:

$$\mu(R_0+1)\mu^2 R_0 + \frac{1}{R_0}\beta\mu(R_0-1) > \frac{1}{R_0}\beta\mu^2(R_0-1).$$

By rearranging the terms in this inequality, we reach to the simplified one:

$$\mu R_0(R_0+1) + \frac{1}{R_0}\beta(R_0-1)(R_0+1) > \frac{1}{R_0}\beta(R_0-1).$$

From $a_3 > 0$, we know that the $R_0$ is greater than 1. There is a term identical in both sides of the inequality $\frac{1}{R_0}\beta(R_0-1)$, which is increased by positive terms in the left hand-side. Meaning that the inequality holds true for $R_0 > 1$. Thus we conclude that the endemic equilibrium is stable for $R_0 > 1$ if and only if $\beta > \mu + \gamma$.

Summarizing the stability properties of the DFE and EE of the model with constant recruitment rate, the above analysis proves the following theorem.

**Theorem 4.1.1** *All solutions of model (4.3) with non-negative initial conditions and a constant recruitment rate $f(N) = \Lambda$ are non-negative. The model has two equilibrium points: disease free equilibrium $(S_0, I_0, R_0) = \left(\frac{\Lambda}{\mu}, 0, 0\right)$ and endemic equilibrium*

$$(S_\infty, I_\infty, R_\infty) = \left(\frac{\Lambda}{\mu R_0}, (R_0 - 1)\frac{\Lambda}{\beta}, (R_0 - 1)\frac{\gamma\Lambda}{\mu\beta}\right),$$

*which exchange stabilities in the following way:*

- *when $R_0 < 1$ then there is only one stable disease-free equilibrium point. An endemic equilibrium is unstable;*
- *when $R_0 > 1$ then there a stable endemic equilibrium point, while the disease free equilibrium point is unstable.*

*The system undergoes a transcritical bifurcation.*

We see that the equilibrium points switch stabilities, causing transcritical bifurcation. Meaning that if a person infected with tuberculosis infects more than one susceptible, then the population quickly gets infected and the epidemic situation in the country becomes endemic. Otherwise, the number of infected people decrease approaching the disease-free state.

The model with constant recruitment rate assumes that the population recruits the same number of new susceptibles each time (in our case, we choose monthly count). As we can see, the stability of disease-free state requires only the number of secondary infections to be less than 1. However, this kind of a model does not fully depict the real growth rate of the population. Hence, there is a ground to try another recruitment rate, namely, proportional.

## 4.2 SIR with Proportional Recruitment Rate

### 4.2.1 Model Formulation

The model here involves the proportional recruitment rate, $r$ meaning the birth rate. Here, the number of new susceptibles entering the population is proportional to its size N.

$$\begin{cases} S'(t) = rN - \frac{\beta SI}{N} - \mu S, \\ I'(t) = \frac{\beta SI}{N} - (\gamma + \mu)I, \\ R'(t) = \gamma I - \mu R. \end{cases} \tag{4.7}$$

The population size changes with time in the following way $N' = (r - \mu)N$. At this point we observe that if $r < \mu$, then $N \to 0$. In other words, if the birth rate is less that the natural death rate then the population size inevitably approaches 0, indicating the extinction of the population. Otherwise, the solutions of this model stay in the biologically reasonable region. Further we consider the case when $r > \mu$. First, the steady states of the system are identified.

### 4.2.2 Equilibria

Following the same technique as in the previous section, we find the equilibrium points, linearize the system around them and look for a characteristic equation. There exist two equilibrium points corresponding to the population extinction $(S_e, I_e, R_e) = (0,0,0)$ and the endemic situation

$$(S_\infty, I_\infty, R_\infty) = \frac{1}{\beta(r - \frac{\mu}{R_0})}\Big((\gamma+\mu)^2, \mu(\beta - \gamma - \mu), \gamma(\beta - \gamma - \mu)\Big).$$

### 4.2.3 Stability

To linearize the model around these equilibrium points, we calculate the Jacobian of the system:

$$J(S,I,R) = \begin{bmatrix} -\frac{\beta I}{N} - \mu & -\frac{\beta S}{N} & 0 \\ \frac{\beta I}{N} & \frac{\beta S}{N} - (\gamma+\mu) & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

Evaluating the matrix at the extinction equilibrium gives the following matrix:

$$J(S_e,I_e,R_e) = \begin{bmatrix} -\mu & 0 & 0 \\ 0 & (\gamma+\mu) & 0 \\ 0 & \gamma & -\mu \end{bmatrix},$$

which we use to find the characteristic equation below:

$$(-\mu-\lambda)(-\gamma-\mu-\lambda)(-\mu-\lambda) = 0. \tag{4.8}$$

The roots of this equation are the eigenvalues we are looking for:

$$\lambda_1 = -\mu, \qquad \lambda_2 = -\mu, \qquad \lambda_3 = -(\gamma+\mu).$$

The population goes extinct if $\mu+\gamma > 0$, meanings this trivial equilibrium is stable. Otherwise, it gets unstable.

If we do the same for the endemic equilibrium, we get the following matrix and the corresponding characteristic equation:

$$J(S_\infty,I_\infty,R_\infty) = \begin{bmatrix} -\frac{\mu(\beta-\gamma-\mu)}{\gamma+\mu} & -(\gamma+\mu) & 0 \\ \frac{\mu(\beta-\gamma-\mu)}{\gamma+\mu} & 0 & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

$$\lambda(\mu+\lambda)(-\mu(R_0-1)-\lambda)+\mu(\gamma+\mu)(R_0-1)(-\mu-\lambda)=0. \tag{4.9}$$

We rearrange the equation (4.9) to clearly see the coefficients $a_1, a_2, a_3$ to be able to apply the Routh-Hurwitz criteria.

$$\lambda^3+\mu R_0\lambda^2+\mu(R_0-1)(2\mu+\gamma)\lambda+\mu^2(R_0-1)(\mu+\gamma)=0. \tag{4.10}$$

The corresponding coefficients are:

$$\begin{aligned} a_1 &= \mu R_0, \\ a_2 &= \mu(R_0-1)(2\mu+\gamma), \\ a_3 &= \mu^2(R_0-1)(\mu+\gamma). \end{aligned}$$

First, we verify that $a_1 > 0$. For that the following inequality must hold: $R_0 > 0$. Whereas the positiveness of the coefficient $a_3$ is supported with the condition $R_0 > 1$ and the positiveness of model parameters $\mu, \gamma > 0$. What is remaining is to see if $a_1 a_2 > a_3$:

$$\mu^2 R_0(R_0-1)(2\mu+\gamma) > \mu^2(R_0-1)(\mu+\gamma).$$

By crossing out identical terms from each side, we reach the following form:

$$R_0(\gamma+2\mu) > \gamma+\mu. \tag{4.11}$$

First, we remember that $R_0 > 1$. Next, parameters $\mu$, $\gamma$, $R_0$ are positive and they either sum up or multiply, proving that the terms are always positive. It also can be observed that $(\gamma+2\mu) > \gamma+\mu$. Hence, we conclude that the Routh-Hurwitz conditions hold and prove the local stability of an endemic equilibrium for $R_0 > 1$. These processes are summarized in a theorem below.

**Theorem 4.2.1** *All solutions of model (4.7) with non-negative initial conditions and a propor-*

*tional recruitment rate $f(N) = rN$ are non-negative. The model has two equilibrium points: an extinction equilibrium $(S_e, I_e, R_e) = (0,0,0)$ and an endemic equilibrium $(S_\infty, I_\infty, R_\infty) = \frac{1}{\beta(r-\frac{\mu}{R_0})}\Big((\gamma+\mu)^2, \mu(\beta-\gamma-\mu), \gamma(\beta-\gamma-\mu)\Big)$, which exchange stabilities in the following way:*

- *when $R_0 > 1$, i.e. $\beta > \gamma+\mu$ an endemic equilibrium point becomes the only one stable equilibrium, otherwise it loses its stability ;*
- *when $r < \mu$, the population goes extinct, leaving no steady states; otherwise, the stability of an extinction scenario is only supported by $\mu + \gamma > 0$.*

In summary, the model (4.7) allows both the possibility of the population extinction and the endemic situation in the region under the assumptions listed in the above theorem.

Let us now examine the logistic recruitment rate, which is a more realistic population growth scenario, since it considers the resources needed for living. First, there are plenty of resources for growth, hence the population size grows accordingly. As the resources become limited, the growth slows down and the population size remains near its carrying capacity and rarely reaches it.

## 4.3 SIR with Logistic Recruitment Rate

### 4.3.1 Model Formulation

In this model, we consider a logistic growth of the population, where $K$ is the carrying capacity of the population.

$$\begin{cases} S'(t) = rN(1-\frac{N}{K}) - \frac{\beta SI}{N} - \mu S, \\ I'(t) = \frac{\beta SI}{N} - (\gamma+\mu)I, \\ R'(t) = \gamma I - \mu R. \end{cases} \tag{4.12}$$

The model assumes that the population grows logistically and it depends both on the actual population size $N$ at time $t$ and its carrying capacity $K$. Solutions of the model lie in the region $\{S \geq 0, I \geq 0, R \geq 0, S+I+R \leq K\}$. To verify it, we assume that $S+I+R \geq K$. Looking at the rate of change of $N$,

$$N' = rN(1-\frac{N}{K}) - \mu N,$$

we can observe that $rN(1-\frac{N}{K}) - \mu N \leq -\mu N$, meaning the decreasing population. Hence we only want to consider the region $\{S \geq 0, I \geq 0, R \geq 0, S+I+R \leq K\}$. In addition, we require $r > \mu$ for the population to grow. Let us now determine its equilibrium points.

### 4.3.2 Equilibria

There are three steady states of the model (4.12). The first one is $(S_e, I_e, R_e) = (0,0,0)$ which corresponds to the extinction scenario, meaning there is no population, hence no disease at all. The next two are the disease-free equilibrium and an endemic equilibrium presented below.

- DFE:

$$(S_0, I_0, R_0) = \left(\frac{(r-\mu)K}{r}, 0, 0\right),$$

- EE:

$$(S_\infty, I_\infty, R_\infty) = \frac{(r-\mu)K}{r}\left(\frac{1}{R_0}, \frac{\mu(R_0-1)}{\beta}, \frac{\gamma(R_0-1)}{\beta}\right).$$

In comparison with the previous models, now we can see that the endemic equilibrium point also depends on the carrying capacity of the population. Following that, we examine the stability properties of these equilibrium points.

### 4.3.3 Stability

The Jacobian matrix evaluated at the population extinction equilibrium is

$$J(S_e, I_e, R_e) = \begin{bmatrix} -\mu & 0 & 0 \\ 0 & -(\gamma+\mu) & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

The characteristic equation obtained from this matrix is

$$(-\mu-\lambda)(-(\gamma+\mu)-\lambda)(-\mu-\lambda) = 0. \tag{4.13}$$

The eigenvalues are immediate:

$$\lambda_1 = \lambda_2 = -\mu, \qquad \lambda_3 = -(\gamma+\mu).$$

Since we know that $\mu > 0$, we look at the third eigenvalue. If $\gamma+\mu > 0$, then the equilibrium will be stable, otherwise unstable. However, we also know that the requirement for population extinction is $r < \mu$. Hence, the former condition is not the only one leading to stable extinction equilibrium.

Considering the disease-free equilibrium, we get to the following Jacobian matrix:

$$J(S_0, I_0, R_0) = \begin{bmatrix} -\mu & -\beta & 0 \\ 0 & \beta-(\gamma+\mu) & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

The characteristic equation obtained from this matrix is

$$(-\mu-\lambda)(\beta-(\gamma+\mu)-\lambda)(-\mu-\lambda) = 0. \tag{4.14}$$

The eigenvalues of the equation can be clearly seen to be

$$\lambda_1 = \lambda_2 = -\mu, \qquad \lambda_3 = \beta - (\gamma + \mu).$$

We observe that the solutions will approach the disease-free equilibrium when $\beta < \gamma + \mu$. Remembering that the basic reproduction number is $R_0 = \frac{\beta}{\gamma+\mu}$, the stability of the DFE is supported by $R_0 < 1$.

Remaining equilibrium point is the endemic one:

$$(S_\infty, I_\infty, R_\infty) = \frac{(r-\mu)K}{r}\left(\frac{1}{R_0}, \frac{\mu(R_0-1)}{\beta}, \frac{\gamma(R_0-1)}{\beta}\right).$$

By evaluating the Jacobian matrix at this point, we get the following matrix:

$$J(S_\infty, I_\infty, R_\infty) = \begin{bmatrix} -\mu R_0 & -\frac{\beta}{R_0} & 0 \\ \mu(R_0-1) & 0 & 0 \\ 0 & \gamma & -\mu \end{bmatrix}.$$

The characteristic equation of this matrix is

$$-\lambda(-\mu R_0 - \lambda)(-\mu - \lambda) + \frac{\beta\mu}{R_0}(R_0-1)(-\mu-\lambda) = 0. \tag{4.15}$$

The equation (4.15) is of third order, hence we consider applying Routh-Hurwitz conditions to check the stability of EE. By bringing this equation into the following form, we clearly see the coefficients of interest:

$$\lambda^3 + (R_0+1)\mu\lambda^2 + \left(\mu^2 + \mu\beta\left(1 - \frac{1}{R_0}\right)\right)\lambda + \beta\mu^2\left(1 - \frac{1}{R_0}\right) = 0. \tag{4.16}$$

The coefficients needed are listed below:

$$
\begin{aligned}
a_1 &= (R_0+1)\mu, \\
a_2 &= \mu^2 + \mu\beta\left(1-\frac{1}{R_0}\right), \\
a_3 &= \beta\mu^2\left(1-\frac{1}{R_0}\right).
\end{aligned}
$$

The first coefficient $a_1$ is positive since $R_0 > 0$ and $\mu > 0$. For coefficient $a_3$ to be positive, $R_0$ has to be greater that 1. To examine the last condition $a_1a_2 > a_3$, we set the inequality and then reformulate it. After rearranging we get the following inequality:

$$
(R_0+1)\left(\mu+\beta\left(1-\frac{1}{R_0}\right)\right) > \beta\left(1-\frac{1}{R_0}\right) \tag{4.17}
$$

It can be observed that there is a term identical on both sides of the inequality $\beta\left(1-\frac{1}{R_0}\right)$. The other parameters on the left hand side are only added or multiplied to this term, hence proving that the expression on the left is only increasing. The inequality is verified, thus proving that the Routh-Hurwitz conditions are satisfied. Hence, if $R_0 > 1$, the endemic equilibrium is stable, otherwise unstable. These results are summarized in the theorem below.

**Theorem 4.3.1** *All solutions of model (4.12) with non-negative initial conditions and a logistic recruitment rate $f(N) = rN(1-\frac{N}{K})$ are non-negative. The model has three equilibrium points: extinction $(S_e, I_e, R_e) = (0,0,0)$, disease free $(S_0, I_0, R_0) = \left(\frac{(r-\mu)K}{r}, 0, 0\right)$ and endemic equilibrium $(S_\infty, I_\infty, R_\infty) = \frac{(r-\mu)K}{r}\left(\frac{1}{R_0}, \frac{\mu(R_0-1)}{\beta}, \frac{\gamma(R_0-1)}{\beta}\right)$, which exchange stabilities in the following way:*

- *when $r < \mu$ and $\mu + \gamma > 0$ then the population goes extinct, and $(S_e, I_e, R_e)$ is the only equilibrium. When $r > \mu$ there two more equilibrium points: DFE and EE;*
- *when $R_0 < 1$, the disease free state is stable, and endemic state is unstable;*
- *when $R_0 > 1$, the endemic equilibrium is stable, and the disease-free state loses its stabil-*

*ity.*

*There is a transcritical bifurcation between the last two equilibrium points when* $r > \mu$.

The logistic growth of the population is the most realistic one due to consideration of the carrying capacity of the population size. It also considers three different possible cases for the disease dynamics, and gives a broader knowledge on this matter.

Summarizing the results of these three sections, tables and some conclusions were drawn in the next section.

## 4.4 Summary

In this section we summarize the equilibrium points and stability properties of all three models in tables. Table 1 represents the equilibrium points.

Table 1: Equilibrium points of SIR models

| Recruitment type | Equilibrium points | | |
|---|---|---|---|
| | Disease-free | Endemic | Extinction |
| Constant | $\left(\frac{\Lambda}{\mu},0,0\right)$ | $\left(\frac{\Lambda}{\mu R_0},(R_0-1)\frac{\Lambda}{\beta},(R_0-1)\frac{\gamma\Lambda}{\mu\beta}\right)$ | |
| Proportional | | $\frac{1}{\beta(r-\frac{\mu}{R_0})}\left((\gamma+\mu)^2,\mu(\beta-\gamma-\mu),\gamma(\beta-\gamma-\mu)\right)$ | $(0,0,0)$ |
| Logistic | $\left(\frac{(r-\mu)K}{r},0,0\right)$ | $\frac{(r-\mu)K}{r}\left(\frac{1}{R_0},\frac{\mu(R_0-1)}{\beta},\frac{\gamma(R_0-1)}{\beta}\right)$ | $(0,0,0)$ |

It is observed that while the model with constant recruitment rate gives the disease-free and endemic equilibrium points, and the proportional growth rate gives only endemic and extinction equilibria, the logistic one provides all of the three possible scenarios.

Now, it is important to summarize stability properties of these equilibria. It is presented in table 2. Remembering from section 4.1, that the basic reproduction number for all of the models is

$$R_0 = \frac{\beta}{\mu+\gamma}$$

, it can be observed that the stability of equilibrium points mainly depends on the $R_0$: whether it is greater or less than 1. An endemic equilibrium is only stable if $R_0 > 1$, while otherwise the

disease-free state is stable. An the population extinction equilibrium is stable when the birth rate is less that the natural death rate $r < \mu$.

Table 2: Stability of equilibrium points

| Recruitment type | Parameter conditions | Epidemic outcome |
|---|---|---|
| Constant | $\beta < \mu + \gamma$ | Disease free |
| | $\beta > \mu + \gamma$ | Epidemic |
| Proportional | $r < \mu$ | Extinction |
| | $\beta > \mu + \gamma$ | Epidemic |
| Logistic | $\beta < \mu + \gamma$ | Disease free |
| | $\beta > \mu + \gamma$ | Epidemic |
| | $r < \mu$ | Extinction |

While both systems (4.3), (4.12) undergo the transcritical bifurcation, there is no such process observed in the case with proportional growth. It is the fact that if $r > \mu$, there is only an endemic equilibrium which is stable if $R_0 > 1$. Whereas the condition $r < \mu$ leads to the extinction of the population according to models (4.7) and (4.12).

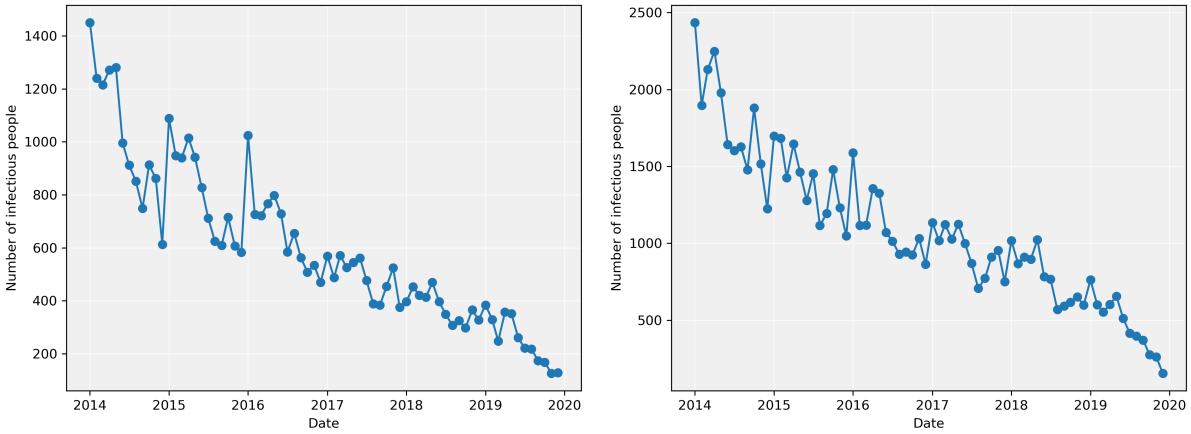
In the next chapter we represent the results of simulations of the three models using the real life data containing the number of monthly TB cases in Kazakhstan over the period from 2014 to 2017.

# Chapter Five

# Tuberculosis Dynamics Simulation and Error Analysis

## 5.1 Simulations

The models of interest were simulated using the data retrieved from the Unified National Electronic Health System (UNEHS). It contains monthly TB cases between January 2014 and December 2019. Only the patients recorded in accordance with the International Classification of Diseases-10 (ICD-10) under the classes A15 (respiratory tuberculosis confirmed bacteriologically and histologically) and A16 (respiratory tuberculosis, not confirmed bacteriologically or histologically) were included into the dataset. The dataset contains only the disease classes, time data and the number of monthly cases. Figure 1 represents the given monthly data. As it can be observed, the dynamics of tuberculosis infection in the country is illustrating a downward trend along with some fluctuations.




(a) A15 - Respiratory tuberculosis, bacteriologically and histologically confirmed.

(b) A16 - Respiratory tuberculosis, not confirmed bacteriologically or histologically.

Figure 1: Monthly data of TB cases in the period 2014-2019

Model parameters were fitted to the data using Python version 3.9.18 with packages for data fitting and optimization was used. To be more precise, a powerful function curve_fit from SciPy library that fits the data by nonlinear least squares method was employed. The dynamics of the disease was simulated using estimated parameters. Although the fit is presented for the data over the whole period of 2014-2019, simulations were done for the first 4 years and the remaining two year data was used to evaluate the performance of the models. Table 3

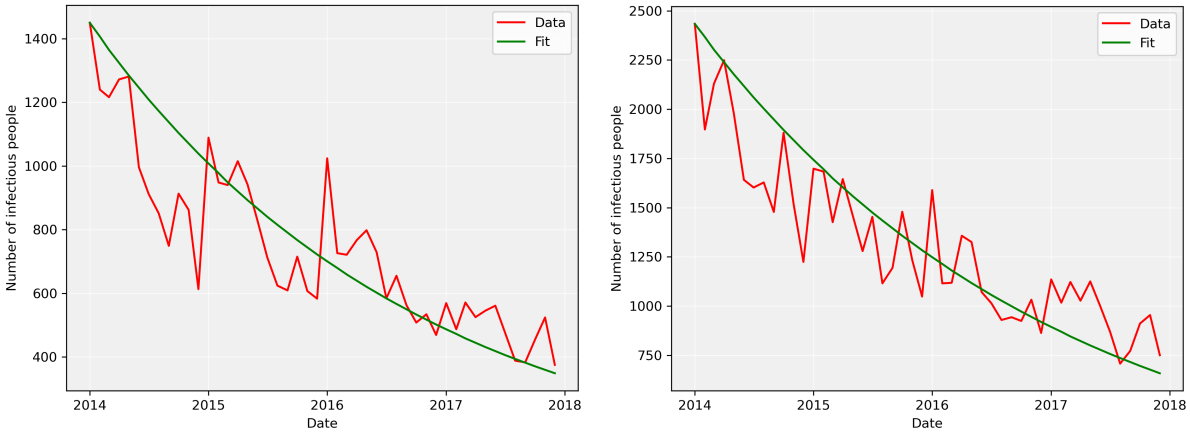
summarizes the values of parameters used for simulations. The results are presented in the following subsections (5.1-5.3). As the epidemic situation is simulated, we wish to evaluate its performance using statistical metrics such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and the determination coefficient (R-squared). First three metrics are the most commonly used ones to evaluate the performance of the model. R-squared, although being mostly used for regression model evaluation, is able to provide the general understanding of how well the model can fit to the data. This combination of metrics is able to give more insights into the model's performance.

Table 3: Parameters of model simulation

| Parameter | Description | A15 class | A16 class | Source |
|---|---|---|---|---|
| $\Lambda$ | Constant recruitment | 22 980/N | 22 980/N | Calculated [42] |
| K | Population carrying capacity | 30 000 000 | 30 000 000 | Assumed |
| $r$ | Monthly recruitment rate | 0.0013 | 0.0013 | Calculated [42] |
| $\gamma$ | Monthly all-cause death rate | 0.0007 | 0.0007 | Calculated [42] |
| $N(0)$ | Initial population size | 17 592 298 | 17 592 298 | [42] |
| $S(0)$ | Initial number of susceptible people | 17 590 848 | 17 589 864 | Calculated |
| $I(0)$ | Initial number of infected people | 1450 | 2434 | Data |
| $R(0)$ | Initial number of recovered people | 1300.36 | 2182.81 | Calculated[20] |

All the computations and plotting were performed using the programming language Python version 3.9.18 with packages for data fitting and optimization. To be more precise, a powerful function curve_fit from SciPy library that fits the data by nonlinear least squares method was employed. Statistical metrics were calculated using the packages from scikit-learn library. All of the three models give the same fitting curve presented in figure 2.

Figure 2 represents the nonlinear curve trying to capture the general behavior of the TB transmission dynamics. In the next sections, we will take a closer look into the simulated dynamics of three compartments of the model, and evaluate how good the simulations are and what effect has the recruitment rate on the epidemic situation in Kazakhstan.



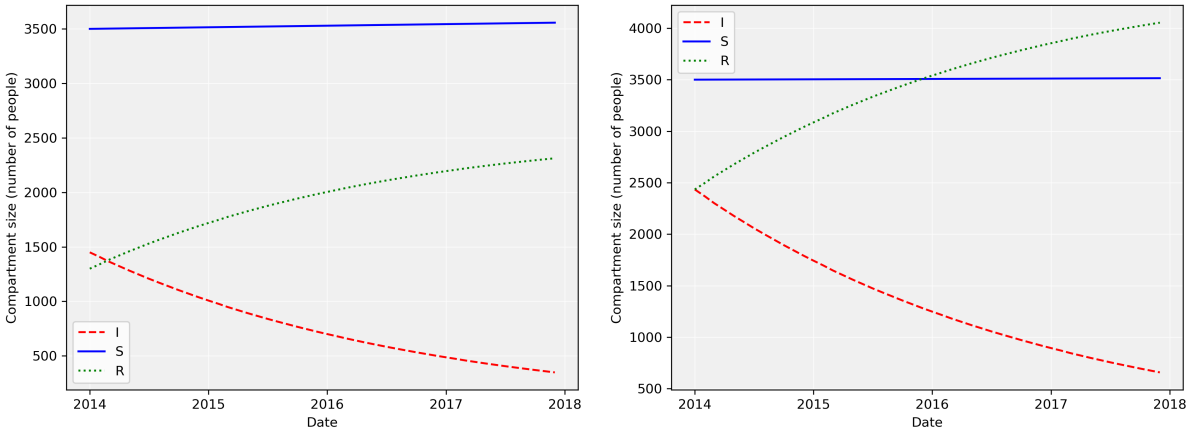

(a) A15 - Respiratory tuberculosis, bacteriologically and histologically confirmed.

(b) A16 - Respiratory tuberculosis, not confirmed bacteriologically or histologically.

Figure 2: Results of model fitting to TB data over the period 2014-2017

### 5.1.1 SIR with constant recruitment rate

In this section, the results of simulating the model (4.3) are presented along with the model's performance evaluation in terms of statistical metrics. Figure 3 illustrates the results of the simulation for both A15 and A16 classes of TB patients.




(a) A15 - Respiratory tuberculosis, bacteriologically and histologically confirmed.

(b) A16 - Respiratory tuberculosis, not confirmed bacteriologically or histologically.

Figure 3: Results of model 1 simulation in the period 2014-2017

As it can be seen from the figure 3, the plot was scaled to fit the sizes of the figure due to a significant difference between the actual number of people in the susceptible and infected

compartments.

In addition, figure 3 illustrates the flow of individuals between compartments. The group of susceptible people illustrates a slight growth in figure 3 (a), while showing no or little action in figure 3 (b). Whereas there are almost symmetrical movements in the recovered and infected groups for both classes of patients.

As the result of fitting and performance analysis, values of contact rate $\beta$, recovery rate $\gamma$, the basic reproduction number $R_0$ , and the statistical metrics were calculated. The results are presented in table 4.

Table 4: Fitting results and error analysis: constant recruitment rate $f(N) = \Lambda$

| ICD-10 Code | A15-confirmed | A16-not confirmed |
|---|---|---|
| Model parameters | | |
| Contact rate $\beta$ | 0.0014 | 0.0021 |
| Recovery rate $\gamma$ | 0.0310 | 0.0292 |
| The basic reproduction number $R_0$ | 0.0442 | 0.0691 |
| Performance evaluation | | |
| R-squared | 0.61 | 0.42 |
| RMSE | 60.60 | 174.55 |
| MAE | 48.85 | 146.11 |
| MAPE | 0.18 | 0.38 |

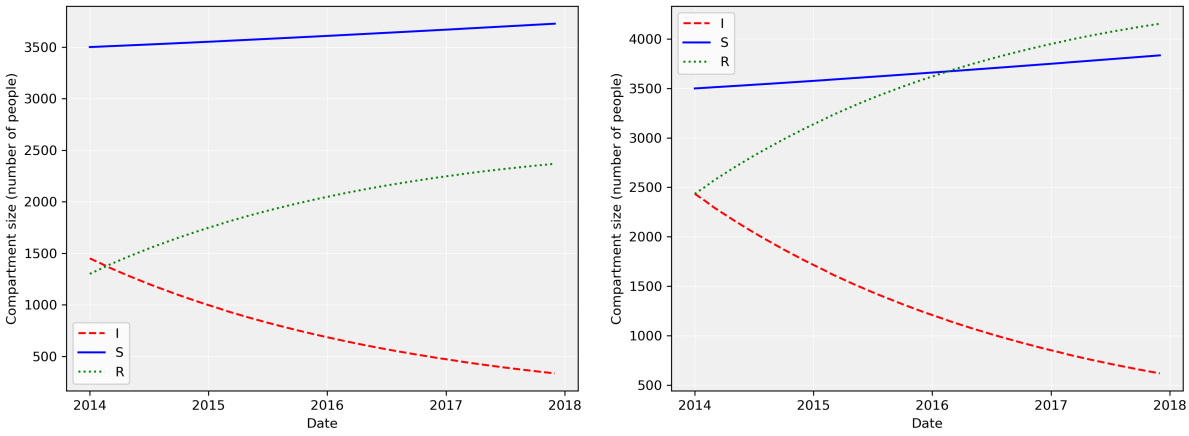
As presented in the table above, fitting the model to data revealed that all of the three parameters $\beta$, $\gamma$, $R_0$ are less than 1, supporting the decreasing trend in TB dynamics in the country (figures 1-3). However, the class A16 shows that the contact rate and the average number of secondary infections is a bit higher than the ones revealed for class A15, while the recovery rates are almost equal, both being $\gamma \approx 0.03$.

Performance evaluation illustrates that the model (4.3) performs better the data of class A15, with R-squared showing 0.61 and MAPE giving 0.18, while the performance for the class A16 is less accurate: R-squared$= 0.42$ and MAPE $= 0.38$. Being aware of the ideal values the models should aim at, i.e. 1 for R-squared and 0 for MAPE, we see that the values for A15 are closer to the ideal model results than for the class A16. RMSE and MAE values for both classes differ dramatically, with the values for A15 being 60.60 and 48.85 and for A16 being greater

than 100.

### 5.1.2 SIR with proportional recruitment rate

The next system to consider is the model (4.7). It was simulated using the same technique as the model with constant recruitment rate. Simulation of compartment dynamics is presented on figure 4.




(a) A15 - Respiratory tuberculosis, bacteriologically and histologically confirmed.

(b) A16 - Respiratory tuberculosis, not confirmed bacteriologically or histologically.

Figure 4: Results of model 2 simulation in the period 2014-2017

The figure illustrates a more rapid increase in the number of susceptibles, proving the proportional population growth. The dynamics of infected and recovered group of people move down and upwards, respectively, in a symmetrical manner. As the number of people recovered from the disease increases, quantity of TB infected people decreases. The movements are monotone. The parameters that support these dynamics are presented in table 5.

The model with proportional recruitment rate revealed similar results for the data of both classes, giving $\beta \approx 0.002$, $\gamma \approx 0.03$ and the basic reproduction number $R_0$ being around 0.07. It can be observed that the rate at which infected people recover is higher than rate at which individuals get infected.

As a result of performance evaluation, we observe that the metrics for the model with proportional recruitment is similar (almost identical) to the model with constant recruitment rate,

Table 5: Fitting results and error analysis: proportional recruitment rate $f(N) = rN$

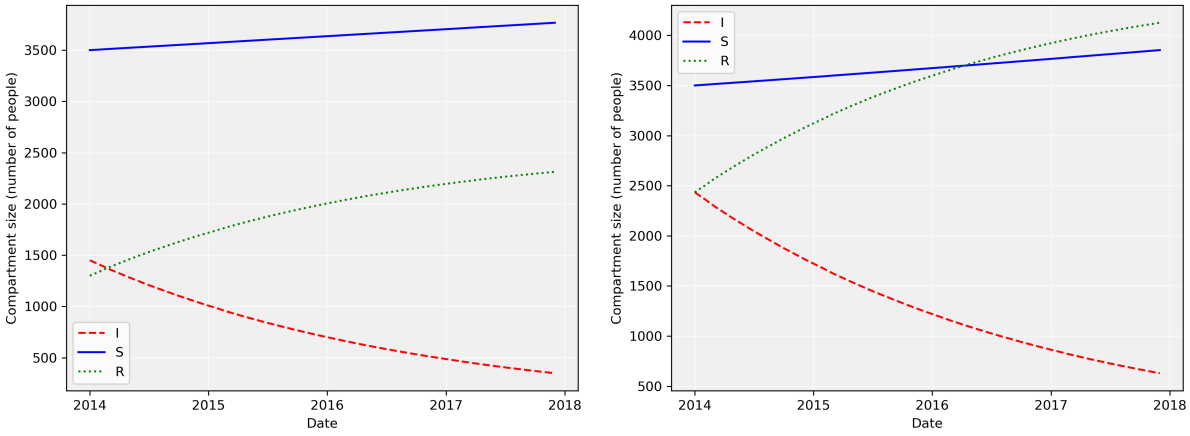
| ICD-10 Code | A15-confirmed | A16-not confirmed |
|---|---|---|
| Model parameters | | |
| Contact rate $\beta$ | 0.0021 | 0.0024 |
| Recovery rate $\gamma$ | 0.0317 | 0.0295 |
| The basic reproduction number $R_0$ | 0.0638 | 0.0788 |
| Performance evaluation | | |
| R-squared | 0.61 | 0.42 |
| RMSE | 60.60 | 174.56 |
| MAE | 48.85 | 146.11 |
| MAPE | 0.18 | 0.38 |

again showing the model worked better with the type of data of class A15.

In the next section, we are willing to examine the performance of the model with logistic growth rate.

### 5.1.3 SIR with logistic recruitment rate

This section presents the results of simulating the model with logistic recruitment rate and examining its ability to describe the dynamics. Simulation is plotted on figure 5.




(a) A15 - Respiratory tuberculosis, bacteriologically and histologically confirmed.

(b) A16 - Respiratory tuberculosis, not confirmed bacteriologically or histologically.

Figure 5: Results of model 3 simulation in the period 2014-2017

The figure depicts almost the same scenario for the model logistic growth as in the case with the one employing proportional growth rate. The susceptible population grows a bit faster,

while the change in the size of infected population influences directly the number of recovered people. It reveals a symmetric movement in both (a) and (b) parts of the figure 5. Parameters used after fitting the model are presented in the table 6.

Table 6: Fitting results and error analysis: logistic recruitment rate $f(N) = rN(1 - \frac{N}{K})$

| ICD-10 Code | A15-confirmed | A16-not confirmed |
|---|---|---|
| Model parameters | | |
| Contact rate $\beta$ | 0.00003 | 0.0017 |
| Recovery rate $\gamma$ | 0.0297 | 0.0288 |
| The basic reproduction number $R_0$ | 0.0011 | 0.0568 |
| Performance evaluation | | |
| R-squared | 0.61 | 0.42 |
| RMSE | 60.60 | 174.56 |
| MAE | 48.85 | 146.11 |
| MAPE | 0.18 | 0.38 |

In this case, parameter values differ notably for the two patient classes. The contact rate estimated for A15 class data 0.00003 is 100 times smaller than the one estimated for the class A16 - 0.0017, indicating that the TB infection of class A15 is more difficult to transfer from one individual to another. Moreover, this is supported by the fact that the number of secondary infections is higher for the A16 class indicating 0.0568, while it is 0.0011 for the class A15. However, the rate at which infected people recover is almost the same.

Performance of the model with logistic growth is identical to the previous models' metrics, as it is presented in tables 4-6. In general, it is clearly seen that all of the model are able to capture some trends in the disease dynamics, but such metrics as RMSE and MAE illustrate that there are uncertainties in their ability to simulate the real-life data.

The next chapter discusses these results in more details along with the theoretical findings.

# Chapter Six

# Discussion

## 6.1 Discussion of Results

In this section we examine the results retrieved from mathematical analysis of models (4.3),(4.7) and (4.12) and simulations along with performance analysis, and discuss the epidemiological situation of tuberculosis infection in Kazakhstan.

As a result of stability analysis of the models, the effect of incorporating various recruitment rates is clearly seen. Referring to table 2 in section 4.4, it is interesting to note that the stabilities of disease-free and endemic equilibrium points depend on the population growth rate and the basic reproduction number. Stability of equilibrium point of the models with proportional and logistic growth rates totally depend on vital dynamics $r$, $\mu$ and the basic reproduction number $R_0$. Zhao et al., who considered a 2D model (SI), also proved that the recruitment rate strongly influences the asymptotic behavior of the model [14]. Their study also reveals three different bifurcation scenarios for each of the HIV models, while this study revealed only the transcritical bifurcation in models (4.3) and (4.12) (constant and logistic recruitment rates) and the change in stability of only the endemic equilibrium of model (4.7) if $R_0 > 1$ and $r > \mu$. The extinction scenario in the latter model is possible if $r < \mu$ even without TB invasion.

It can be observed that the simulation results agree with the theoretical findings. First of all, SIR with proportional and logistic recruitment rates illustrate notable population growth throughout the period of four years, while the model with constant growth represents a little or no change in the population size, since the difference is negligible. This proves that the population size is increasing as time passes, and the solutions of the models lie in the biologically feasible region, as shown in the previous chapter. Next, figures 3-5 illustrate the decreasing trend in the number of TB patients of both classes A15 and A16, while the number of recovered people is increasing. As indicated in theorems 4.1.1 and 4.3.1, the population approaches its disease-free equilibrium if the number of secondary infections is less that 1. According to the results of simulations and calculations, one infected person infects less than one person on average, justifying the dynamics plotted in figures 3-5 and proving that the TB dynamics in

Kazakhstan is approaching its disease-free equilibrium. This can also be supported by successful treatment strategies employed in Kazakhstan [43].

Although acting distinctly, the models were able to simulate the general trend of tuberculosis transmission in Kazakhstan throughout the years from 2014 until 2019. All of the models performed well according to the performance analysis. Tables 4-6 convince us in the identical performance of the models when given the real data on TB cases. However, the models illustrated that the rate at which the disease transmits also depends on the population growth rate. The best scenario for the population is a logistic growth, which showed that the contact rate might be as small as 0.00003 for A15 class and 0.0017 for A16 class of TB patients, while the basic reproduction number was 0.0011 and 0.0568 for A15 and A16 classes respectively (table 6). Whereas the highest possible contact rate and the largest number of possible secondary infections was performed by the model with proportional growth rate (table 5). Similar results were obtained by Ucakan et al. while examining the TB dynamics in Turkey using three models [7]. The difference of this study is in consideration of various recruitment rates, while the one by Ucakan et al. compared various compartments' effect. However, we brought up similar results with $\beta$,$\gamma$ and $R_0$ being less than one and the models capturing main trends. It also should be noted, that the ability of a model with constant growth rate to describe the disease over a short period of time was proven, while the model with proportional and logistic growth rates

According to our results, if the population of the country grows logistically, there are higher chances of TB dying out in future, although it is the ideal case scenario. In addition, looking at the results of all the three models, we may state that the situation with TB in Kazakhstan is under control. However, it might get worse if the contact rate exceeds the natural death rate in summation with the recovery rate; or if the basic reproduction number exceeds 1.

## 6.2 Limitations

This work has got a few limitations, that might be overcome in further studies. First of all, the models employed in this study although being able to capture the general trends of TB

dynamics, are not able to capture possible seasonal fluctuations appearing in the data (figures 1-2). There is a great chance that the fluctuations continue happening and the dynamics drawn in this paper will not be able to fully explain the epidemic situation in the country. In other words, the model lacks realism. Getting the model closer to reality might assist in proposing more efficient control strategies for TB.

Second limitation is the lack of data. The models considered in this study also employ birth and death rates and the number of recovered people. Most of the parameters in table 3 are calculated or assumed relying on the internet resources, although being found on official websites. If the data was retrieved from the official database, calculations and simulations would have been more accurate.

Last but not least, the gaps in the model analysis are assumed to exist. In this work, only the local stability of equilibrium points was considered. However, the global stability would give broader knowledge about the behavior of the disease dynamics.

# Chapter Seven

# Conclusion and Future Work

## 7.1 Conclusion

This thesis studied the tuberculosis dynamics in Kazakhstan employing the basic epidemiological SIR model with three distinct population growth rates: constant, proportional and logistic, and numerical simulations using the data retrieved from the national health system.

The dynamical analysis of the model was performed and it proved that the local stability of disease-free and endemic equilibrium points depends on the basic reproduction number and population growth rate. Stability analysis also illustrated that the system undergoes transcritical bifurcation under conditions specific to each recruitment rate. In addition, performance of the models and theoretical results were verified and evaluated using numerical simulations and error analysis. The study revealed that the best case scenario was offered by the SIR with logistic recruitment rate, that revealed the smallest contact rate and the basic reproduction number. Despite this fact, all three models were able to capture the general trends of tuberculosis dynamics and give satisfactory results on performance evaluation. Scenarios when the disease might get endemic for the country were also discussed. The situation with TB in Kazakhstan was proven to be under control.

## 7.2 Recommendations for Further Research

The idea of the thesis can further be developed in various ways. First of all, since the data presented in this work is fluctuating with noise at some points, stochasticity might be added to the model to capture this trend and to equip the model with the ability to accurately forecast the future dynamics of the disease. Alternatively, a time-dependent or periodic contact rate $\beta(t)$ might be of great use to capture these fluctuations, that might turn out to be seasonal. Secondly, since we only considered local stability of the equilibrium points, we wish to go further and explore its global stability properties. Lastly, it would be useful for the healthcare policies if these models included the effect of control strategies, such as vaccination in the early age, directly observed therapy (DOT) for TB and isolation of an infected person. This could allow

us to assess the efficacy of various control methods and might lead to a new complex treatment method. Moreover, if there are more data on patients' age-group or place of residents, the effect of contact between youngsters and adults, as long as the effect of place of residence on TB dynamics in Kazakhstan might be an interesting subject to study.

# References


[1] J. Gani, "Daniel Bernoulli," *Statisticians of the Centuries*, pp. 64–67, 2001.

[2] Centers for Disease Control and Prevention, "Ross and the Discovery that Mosquitoes Transmit Malaria Parasites," 2015.

[3] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, pp. 700–721, 1927.

[4] Q. Chen, S. Yu, J. Rui, Y. Guo, S.Yang, G. Abudurusuli, Z. Yang, C. Liu, L. Luo, M. Wang, Z. Lei, Q. Zhao, L. Gavotte, Y. Niu, R. Frutos, and T. Chen, "Transmissibility of tuberculosis among students and non-students: an occupational-specific mathematical modelling," *Infectious Diseases of Poverty*, vol. 11, no. 117, 2022.

[5] D. Kereyu and S. Demie, "Transmission dynamics model of tuberculosis with optimal control strategies in haramaya district, ethiopia," *Advances in Difference Equations*, vol. 2021, no. 289, 2021.

[6] D. K. Das, S. Khajanchi, and T. Kar, "The impact of the media awareness and optimal strategy on the prevalence of tuberculosis," *Applied Mathematics and Computation*, vol. 366, no. 124732, 2020.

[7] Y. Ucakan, S. Gulen, and K. Koklu, "Analysing of tuberculosis in Turkey through SIR, SEIR and BSEIR mathematical models," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 27, pp. 179–202, 2021.

[8] M. O. Favorov, M. Ali, A. Tursunbayeva, I. Aitmagambetova, P. E. Kilgore, S. Ismailov, and T. Chorba, "Comparative tuberculosis (TB) prevention effectiveness in children of Bacillus Calmette-Guérin (BCG) vaccines from different sources, Kazakhstan," *PLOS ONE*, vol. 7, no. 3, 2012.

[9] A. Terlikbayeva, S. Hermosilla, S. Galea, N. Schluger, S. Yegeubayeva, T. Abildayev, T. Muminov, F. Akiyanova, L. Bartkowiak, Z. Zhumadilov, A. Sharman, and N. El-Bassel, "The impact of the media awareness and optimal strategy on the prevalence of tuberculosis," *BMC infectious diseases*, vol. 12, no. 262, 2012.

[10] A. M. Aringazina, N. Aitambaeva, L. Nazarova, G. Alimbekova, S. Ismailov, M. Adenov, P. Jazybekova, G. Musabekova, and E. Alikeyeva, "Awareness level of the population and key groups of the republic of kazakhstan in matters of tuberculosis," *Science & Healthcare*, vol. 23, no. 5, pp. 67–77, 2021.

[11] World Health Organization, "Tuberculosis," 2022.

[12] Y. Sakko, M. Madikenova, A. Kim, D. Syssoyev, K. Mussina, A. Gusmanov, G. Zhakhina, S. Yerdessov, Y. Semenova, B. L. Crape, A. Sarria-Santamera, and A. Gaipov, "Epidemiology of tuberculosis in kazakhstan: data from the unified national electronic healthcare system 2014 - 2019," *BMJ Open*, vol. 13, no. 10, 2023.

[13] World Health Organization, "Global tuberculosis report 2023, Geneva," 2023.

[14] Y. Zhao, D. Wood, H. V. Kojouharov, Y. Kuang, and D. Dimitrov, "Impact of population recruitment on the HIV epidemics and the effectiveness of HIV prevention interventions," *Bulletin of Mathematical Biology*, vol. 78, no. 10, pp. 2057–2090, 2016.

[15] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics-I," *Bulletin of Mathematical Biology*, vol. 53, pp. 33–55, 1991.

[16] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity," *Bulletin of Mathematical Biology*, vol. 53, pp. 89–118, 1991.

[17] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics-II," *Bulletin of Mathematical Biology*, vol. 53, pp. 57–87, 1991.

[18] D. J. Earn, "A light introduction to modelling recurrent epidemics," *Lecture Notes in Mathematics*, vol. 1945, 2008.

[19] C. Castillo-Chavez and B. Song, "Dynamical models of tuberculosis and their applications," *Mathematical Biosciences and Engineering*, vol. 1, no. 2, pp. 361–404, 2004.

[20] F. L. Azizan, S. Sathasivam, M. Khan, and M. Ali, "Study of transmission of tuberculosis by sir model using runge-kutta method," *Journal of Quality Measurement and Analysis*, vol. 18, pp. 13–28, 2022.

[21] P. van den Driessche, "Reproduction numbers of infectious disease models," *Infectious Disease Modelling*, vol. 2, pp. 288–303, 2017.

[22] P. van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, pp. 29–48, 2002.

[23] S. H. Strogatz, *Nonlinear dynamics and chaos*. Westview Press, 2015.

[24] Z. Bai, "Threshold dynamics of a time-delayed SEIRS model with pulse vaccination," *Mathematical Biosciences*, vol. 269, pp. 178–185, 2015.

[25] H. W. Hethcote, "Qualitative analyses of communicable disease models," *Mathematical Biosciences*, vol. 28, pp. 335–356, 1976.

[26] F. Brauer, "Compartmental models in epidemiology," *Lecture Notes in Mathematics*, vol. 1945, 2008.

[27] H. E. Soper, "The interpretation of periodicity in disease prevalence," *Journal of the Royal Statistical Society*, vol. 92, pp. 34–73, 1929.

[28] W. H. Frost, "How much control of tuberculosis?," *American Journal of Public Health and the Nations Health*, vol. 27, no. 8, pp. 759–766, 1937.

[29] F. M. Feldmann, "How much control of tuberculosis: 1937-1957-1977?," *American Journal of Public Health and the Nations Health*, vol. 47, no. 10, pp. 1235–1241, 1957.

[30] C. E. Palmer, L. Shaw, and G. W. Comstock, "Community trials of BCG vaccination.," *American Review of Tuberculosis and Pulmonary Diseases*, vol. 77, no. 6, pp. 877–907, 1958.

[31] H. T. Waaler, A. Geser, and S. Andersen, "The use of mathematical models in the study of the epidemiology of tuberculosis," *American Journal of Public Health and the Nations Health*, vol. 52, no. 6, pp. 1002–1013, 1962.

[32] S. Brögger, "Systems analysis in tuberculosis control: a model.," *American Review of Respiratory Disease*, vol. 95, no. 3, pp. 419–34, 1967.

[33] C. ReVelle, W. R. Lynn, and F. M. Feldmann, "Mathematical models for the economic allocation of tuberculosis control activities in developing nations1,2.," *American Review of Respiratory Disease*, vol. 96, no. 5, pp. 893–909, 1967.

[34] M. Bilal, I. Ahmad, S. A. Babar, and K. Shahzad, "State Feedback and Synergetic controllers for tuberculosis in infected population," *IET Systems Biology*, vol. 15, no. 3, pp. 83–92, 2021.

[35] Y. Zhao, M. Li, and S. Yuan, "Analysis of transmission and control of tuberculosis in mainland China, 2005–2016, based on the age-structure mathematical model," *International Journal of Environmental Research and Public Health*, vol. 14, no. 10, 2017.

[36] S.-M. Lee, H. Y. Park, H. Ryu, and J. Kwon, "Age-specific mathematical model for tuberculosis transmission dynamics in South Korea," *Mathematics*, vol. 9, no. 8, p. 804, 2021.

[37] S. K. Kabunga, E. F. D. Goufo, and V. H. Tuong, "Analysis and simulation of a mathematical model of tuberculosis transmission in Democratic Republic of the Congo," *Advances in Difference Equations*, vol. 642, no. 2020, 2020.

[38] S. Side, W. Sanusi, M. K. Aidid, and S. Sidjara, "Global stability of SIR and SEIR model for tuberculosis disease transmission with Lyapunov function method," *Asian Journal of Applied Sciences*, vol. 9, no. 3, pp. 87–96, 2016.

[39] P. Olver and C. Shakiban, *Applied Linear Algebra*. Springer, New York, 2018.

[40] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Academic Press, 2013.

[41] T. C. Sideris, *Ordinary differential equations and dynamical systems*. Atlantis Press, 2013.

[42] United Nations Population Fund, "World population dashboard - Kazakhstan," 2023.

[43] World Health Organization, "Global tuberculosis report 2020, Geneva," 2020.